图书情报工作
第 65 卷 第 15 期 2021 年 8 月

IPC 分类修订中的技术演化研究

■ 许景龙[1] 赵亚娟[2,3]
[1] 中共广东省委党校图书馆 广州 510053 [2] 中国科学院文献情报中心 北京 100190
[3] 中国科学院大学经济管理学院图书情报与档案管理系 北京 100190

摘 要：[目的/意义]从专利分类修订角度研究技术演化，为技术演化研究提供新思路。[方法/过程]首先根据 IPC 分类表 H 部 2009 - 2018 年的修订情况，总结出新增分类、删除分类、类内转移分类、类间转移分类 4 种修订类型。其次针对分类修订后产生的过档文献提出基于 Word2vec + TextCNN 模型的过档文献再分类方法，使新旧版分类表通过再分类专利产生衔接。最后结合 H 部 2009 年 - 2018 年被修订分类及再分类专利进行技术演化初步探索。[结果/结论]专利再分类模型可有效解决过档文献问题，为专利再分类工作提供参考，同时可衔接新旧版专利分类表；结合 IPC 分类修订及再分类专利可分析分类修订中的主要技术演化方向，为技术演化研究提供新视角。

关键词：IPC 分类 专利分类修订 专利再分类 技术演化
分类号：G250
DOI：10.13266/j.issn.0252 - 3116.2021.15.016

技术演化是指对一定时期内技术发展过程，包括技术发展历史、现状以及未来趋势的描述，通过专利文献计量分析、文本挖掘、趋势预测等可实现技术演化描述，从而掌握技术发展轨迹。现有研究中基于专利分析研究技术演化的方法通常可划分为两大类：①基于专利元数据的技术演化研究，例如对专利数量、申请地区、申请人等统计排序或对专利分类共现关系、专利引用关系进行分析，这类研究着眼于分析专利外部著录项随时间推进过程中的趋势或关联关系，如谢寿峰从专利申请量、申请人等角度建立了技术演变的研究框架[1]；刘云等从专利申请量和增长率的二维矩阵、国家/地区、机构等结合专利质量分析全球碳纳米管的技术创新特征[2]；J. Suzuki 等使用 IPC 主分类、扩展分类以及 JPO 特有分类"facet code"分析了技术演化轨迹[3]。②基于专利内容的技术演化研究，常通过自然语言处理、文本挖掘等技术手段进行主题内容和主题强度的演化分析。如廖列法等结合 LDA 主题模型与 IPC 分类来分析主题强度、主题内容和技术主题强度[4]；陈伟等结合 LDA 模型和 HMM 模型分析技术主题演化[5]；J. Yoon 等基于 SAO 语义结构分析技术趋势[6]；H. Sasaki 等通过时间序列上的 IPC 共现研究技术融合，识别技术机会[7]。以上两类研究中，既有将专利分类作为元数据开展的技术演化研究，也有将专利分类作为技术内容开展的技术演化研究。

将专利分类作为元数据的技术演化研究常从分类号的统计和共现角度出发，通过分类数量、频次、共现关系网络的特征等分析技术现状和趋势。如 J. Zheng 等通过统计 6 个行业专利的 USPC 号分析中国 2003 - 2008 年的工业和技术发展[8]；X. Zhou 等通过专利分类号追踪技术发展路径[9]；S. Jun 通过挖掘专利分类的关联规则判断技术趋势[10]；黄斌等基于专利分类共现识别技术关联特征[11]；J. K. C. Chen 等通过专利和专利分类复合网络研究日本、美国等国家或地区的太阳能电池技术演化[12]；李瑞茜、J. Krafft 等使用分类共现网络的中心度、结构洞和中间人方法识别技术关联网络中的核心技术、中介技术和新兴技术[13-14]。此外也多见将专利分类作为技术内容开展的技术主题演化。如方曙等用分类号代替关键词，对基于语义的分类号 - 文档矩阵进行聚类，从而研究技术演化[15]；吴红等将 IPC 分类作为技术语境提出 WI-LDA 模型，以改进传统 LDA 模型主题辨识性低、可解释性弱等问题[16]；雷滔、苗红等以主 IPC 和追加 IPC 的流向作为知识流向来研究技术跨领域关联度特征[17-18]等。

目前无论将专利分类作为元数据还是技术内容开

作者简介：许景龙（ORCID：0000 - 0002 - 0067 - 7359），助理馆员；赵亚娟（ORCID：0000 - 0003 - 3501 - 8131），研究员，博士，博士生导师，通讯作者，E-mail：zhaoyj@mail.las.ac.cn。

收稿日期：2021 - 02 - 04 修回日期：2021 - 05 - 10 本文起止页码：140 - 152 本文责任编辑：易飞

展技术演化研究,大多直接将专利分类体系作为既成知识体系使用,很少考虑到专利分类体系也随技术演化而被修订,因此从专利分类修订角度探索技术演化的研究成果也较少。F. Lafond 等从专利分类数量、专利分类规模、重新分类 3 个方面研究 USPC 的长期动态[19];兰德公司根据新增专利分类的时间来确定专利权人在技术领域中的领导型或跟随型地位[20];CC. Wang 等基于美国专利重新分类的 5 种类型来研究技术演化[21];王文晶讨论了新能源汽车产业关键技术与专利分类修订的关联[22];牛力等从空调领域的 IPC 修订描述了技术变化[23],这是少有的从专利分类修订的角度关注技术演化的研究。

本文从 IPC 分类修订出发,提出基于 IPC 分类修订的技术演化研究这一命题,并设计专利再分类模型实现专利自动再分类,用再分类专利联系新旧版 IPC 分类体系,意在结合被修订分类及再分类专利描述技术演化情况,通过对比新旧版分类表的修订结果来表示技术演化结果,用被修订分类的专利由旧版分类转移到新版分类的过程亦即过档文献的再分类过程来表示技术演化过程。本文通过 IPC 分类修订和专利再分类实现技术演化的研究一方面可以丰富技术演化研究方法体系,为技术演化研究提供新视角,另一方面本文提出的专利再分类方法为基于 IPC 分类实现技术演化研究提供了实现方法和路径,可为世界知识产权组织的专利再分类工作提供思路,从而完善 IPC 分类作为专利文献检索和管理工具的职能。

1 IPC 分类及其修订

IPC 分类是根据 1971 年签订的斯特拉斯堡协定编制[24],世界知识产权组织(WIPO)认为"为保持 IPC 分类作为专利文献检索和管理的有效工具,它必须是动态的,需要不断的修订"[25],目前 IPC 分类已更新至第八版,2006 年开始使用的第八版迄今也已修订十余次,2009 开始 WIPO 固定在每年 1 月 1 日公布新版分类表。IPC 分类的历史版本及各版的有效时间见表 1。

IPC 分类的修订内容包括分类号、分类标题(官方注解)、索引、附注和参见。其中分类号和分类标题是决定技术范围的信息;索引、附注和参见是使用 IPC 分类表的指引性信息。但索引、附注和参见的修订不涉及技术变化,因此本文只关注分类号和分类标题的修订,基于此总结了 4 种 IPC 分类修订类型:新增分类、删除分类、类内转移分类和类间转移分类。

表 1 各版 IPC 分类表及有效时间

IPC 分类表版本	启用时间	停用时间	更新周期
第一版	1968.9.1	1974.6.30	5 年
第二版	1974.7.1	1979.12.31	5 年
第三版	1980.1.1	1984.12.31	5 年
第四版	1985.1.1	1989.12.31	5 年
第五版	1990.1.1	1994.12.31	5 年
第六版	1995.1.1	1999.12.31	5 年
第七版	2000.1.1	2005.12.31	5 年
第八版	2006.1.1	至今	1 年

注:资料来源于 WIPO

(1)新增分类是指直接对原有分类号进行细分,由于技术发展,原有分类表示的技术主题能以更加精细的方式进行细分时,就需要新增分类,如图 1 所示 H01L33/00 大组的修订即为新增分类。

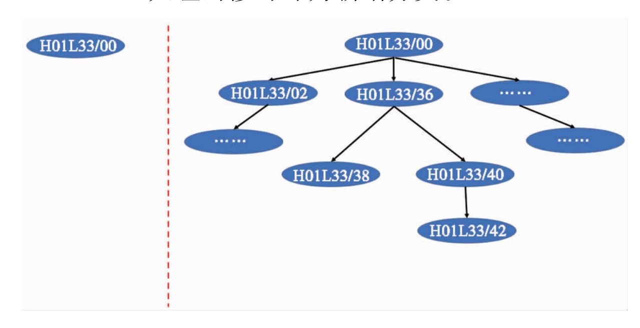


图 1 H01L33/00 大组的新增模式示意

(2)删除分类是指在新版分类表中删除旧版分类号,通常会将被删除分类迁移至分类表中其他位置。而分类的迁移结果又表现为两种情况:一是分类号直接迁移至其他分类,即将该分类与其他分类合并;二是分类号改变,但包含的技术主题并未改变。

(3)类内转移分类。类内转移分类是指分类号在类内进行迁移,判断标准是原分类号被删除,将该技术主题转移至新分类号,并且新分类号与原分类号的上位类相同,如图 2 所示 H04H 小类修订即为类内转移。

(4)类间转移分类。类间转移分类是指分类号在类间迁移,判断标准是原分类号被删除并转移至新分类,新分类与原分类的上位类不同,如图 3 所示 H02S 小类的修订即为类间转移。

从世界知识产权组织目前的工作流程和机制来看,专利分类修订后并不对已分类专利进行重新分类,因此产生"过档文献",即由于 IPC 分类被修订,已分类专利并不随分类修订而调整分类号,因此当使用新版分类号进行专利检索时,使用旧分类号的专利文献无法作为检索结果反馈给用户,这部分未随专利分类修订而调整分类号的专利文献就成为过档文献。过档文

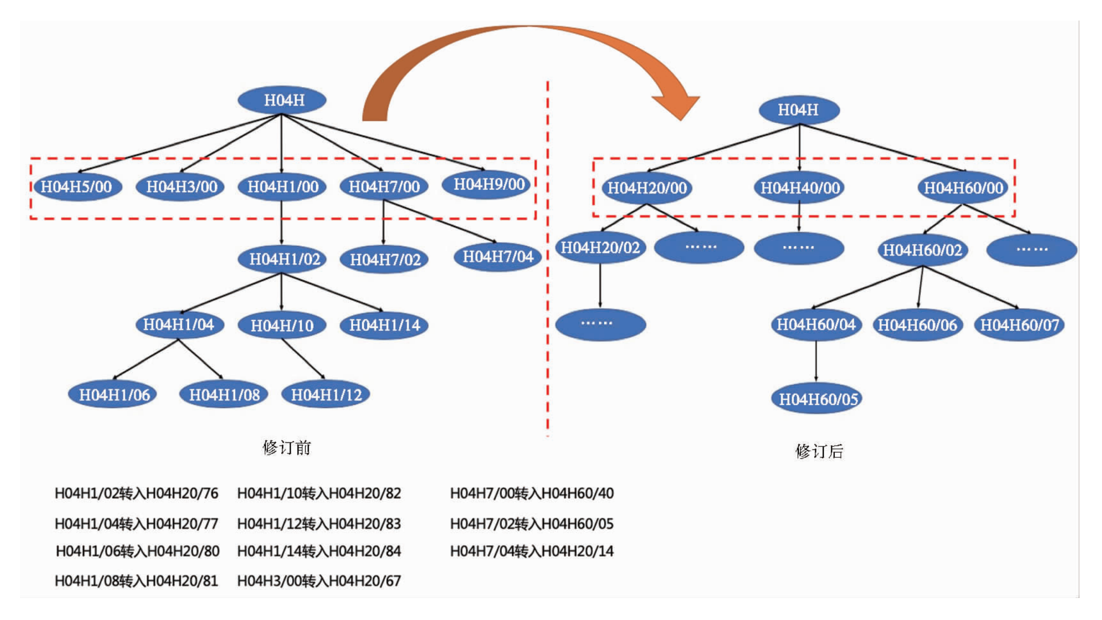


图 2　H04H 小类修订的类内转移模式示意

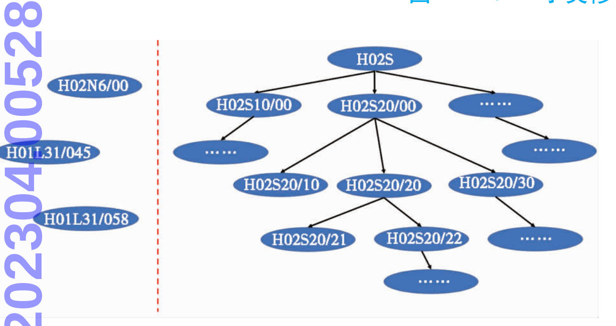


图 3　H02S 小类的类间转移模式示意

献使得基于 IPC 分类修订研究技术演化时存在不同分类版本之间的不连续问题，即旧版分类和新版分类在修订前后无法衔接，因此仅依靠修订前后两个版本分类表只能观察技术演化结果但无法描述技术演化过程。鉴于此，本文设计了基于 Word2vec + TextCNN 模型的过档专利文献再分类方法，通过对过档文献再分类建立新旧版专利分类间的映射关系，并利用再分类专利来描述分类修订后的专利流向信息，从而描述技术演化过程。

2　数据来源与研究方法

2.1　数据来源

由上述介绍可知第八版 IPC 分类表是目前最新版分类表，并且自 2009 年开始统一在每年 1 月 1 日更新新版本，因此本文选择 2009 年 – 2018 年作为研究的时间窗。另外，比较 IPC 分类表中 A – H8 个部在 2009 年 – 2018 年的修订次数，H 部共修订 20 次，是修订次数最多的部，表明 H 部（电学）领域的技术创新和发展速度较快，使用 H 部数据做案例研究更容易得出技术演化的特征。因此本文选择第八版 IPC 分类表 H 部被修订分类作为切入点，以 2009 年 – 2018 年被修订分类作为检索限制条件，以该时间段 H 部被修订分类的公开专利作为数据源进行研究。

2.2　专利再分类模型

专利再分类实际是对过档文献进行文本自动分类，本文提出基于 Word2vec + TextCNN 模型的专利再分类方法，通过机器学习模型解决专利分类修订后的过档文献问题。

Word2vec 是 2013 年 Google 开源的一款词向量（word embedding）计算工具，是词嵌入（word embedding）的一种，也是一种无监督的学习模型，能将词语嵌入到一个数学空间[26]。简单地说，Word2vec 可以在给定的语料库上进行训练，得到预训练模型，使用该模型能输出所有出现在上述语料库中单词的向量表示，基于这个向量可以计算词与词之间的关系，如词与词之间的相似性、语义关联关系等。

TextCNN 模型是 2014 年由 Y. Kim 在 Convolutional Neural Networks for Sentence Classification 一文中提出[27]，该算法利用卷积神经网络 CNN 进行文本分类任务，利用多个不同 size 的 kernel 来提取句子中的关键信息，从而能够更好地捕捉局部相关性。

本文提出基于 Word2vec + TextCNN 模型的专利再分类方法，主要包括数据检索、数据获取、数据预处理、文本分词、Word2vec 模型训练、专利文本向量化、TextCNN 模型训练、专利再分类等过程，完整的技术实现流程如图 4 所示：

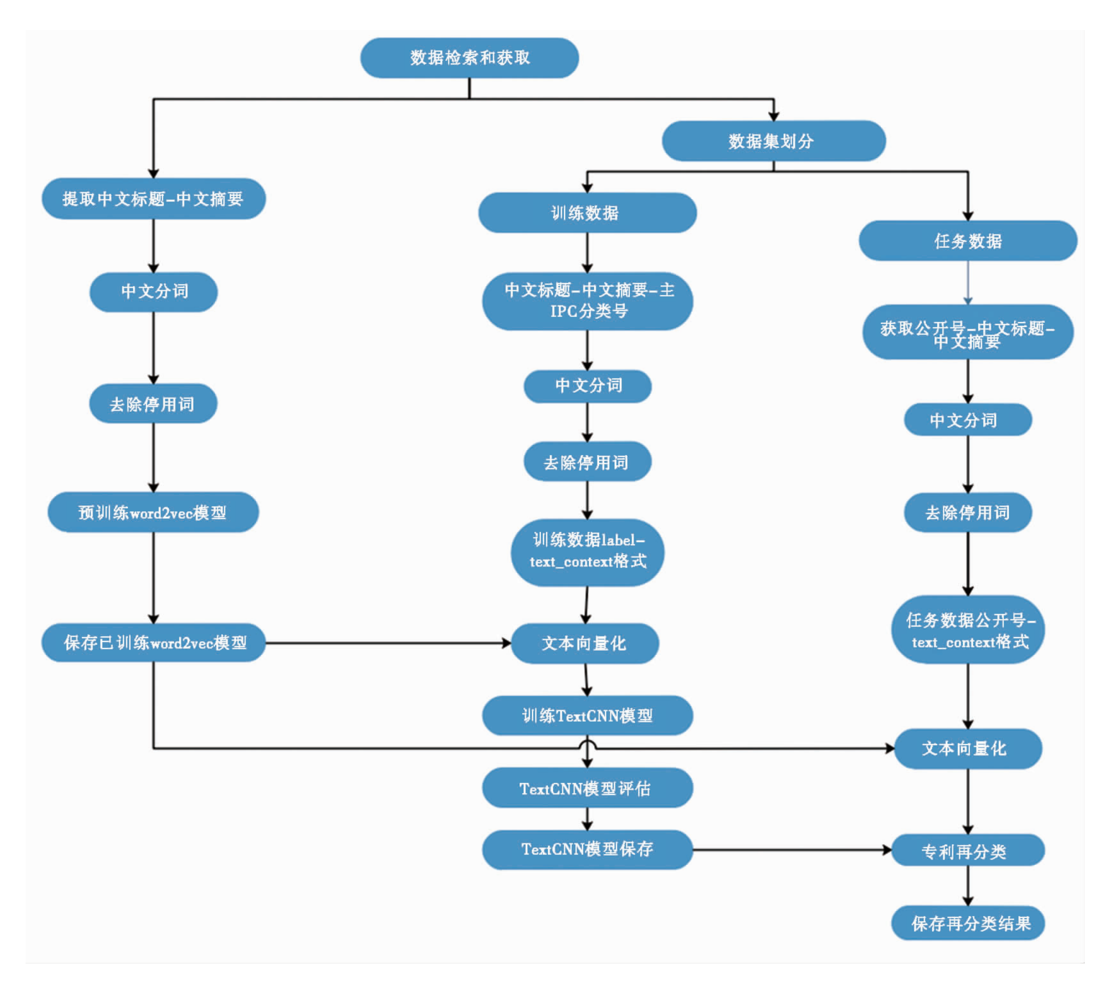


图 4　基于 Word2vec + TextCNN 模型的过档文献再分类技术实现流程

(1)数据检索和获取。以 WIPO 公布的修订分类作为检索条件,选择 Incopat 专利数据库作为数据来源,Incopat 专利数据库具备自动翻译功能,可将外文专利翻译为中文,便于后续中文专利文本处理,因此在 Incopat 专利数据库中检索并获取专利文献集合。

(2)中文分词并去除停用词。选择 jieba 分词工具进行中文分词,去除如标点符号、特殊字符、语气词等停用词,将去除停用词后的中文标题和摘要作为 Word2vec 模型的训练语料集。

(3)训练 Word2vec 模型。本文借用 python 中 NLP 包 gensim 训练 Word2vec,模型输入语料为(2)中已分词的专利标题和摘要,保存训练后的 Word2vec 模型。

(4)数据集划分。以 IPC 分类修订时间为界限,修订时间之前公开的专利为过档专利文献,也是再分类的任务数据集;修订时间点之后公开的专利为训练数据集。

(5)训练数据和任务数据分词、去除停用词。分别对训练数据和任务数据进行分词、去除停用词操作。

(6)文本向量化。调用(3)中已训练的 Word2vec 模型,计算训练数据和任务数据的专利文本向量,作为 TextCNN 模型的训练语料。

(7)TextCNN 模型训练。借助 tensorflow 工具构建 TextCNN 模型并将(6)中已分词的训练语料作为模型输入,训练专利再分类模型。在训练过程中随机选择 80% 的训练数据作为生成模型参数的训练集,另外 20% 的训练数据作为模型效果的测试集,通过测试集数据的分类结果计算模型准确率,选择此种方式计算模型准确率的原因在于目前世界知识产权组织未对过档文献进行再分类,所以无法获取到专业分类员标注数据(分类员手工再分类专利数据)作为模型性能验证数据集。

(8)使用已训练再分类模型对(4)中的任务数据集(过档文献)进行再分类,并保存再分类结果,见表 2 示例。

2.3　模型性能评价

再分类模型性能评价可借鉴常用机器学习模型评价指标,如准确率和损失函数。准确率(Accuracy)表示样本被正确分类的比例,直接用来衡量文本分类模型的效果,通常准确率越高,说明模型分类效果越好。

表 2 过档文献再分类结果示例

公开号	修订前	修订后	公开号	修订前	修订后
US10529868B2	H02N6/00	H02S40/34	AP3824A	H02N6/00	H02S40/30
US10529871B2	H02N6/00	H02S40/22	BRPI1314051A2	H02N6/00	H02S20/00
BRPI1202097A2	H02N6/00	H02S20/00	IN8735DELNP2011A	H02N6/00	H02S40/22
US10361325B2	H02N6/00	H02S40/36	WO2016057341A1	H02N6/00	H02S40/38
US20180254368A9	H01L31/058	H02S50/00	WO2015157328A3	H01L31/058	H02S10/30

$$准确率(Accuracy, acc) = \frac{分类正确的样本数量}{总样本数量}$$

损失函数(Loss)用来计算分类后预测值与真实值之间的距离大小,当预测值与真实值的差距越大时,损失函数 Loss 越大,则需要优化模型减小 Loss;损失函数越小时,说明模型分类效果越好。本文中使用的是 L2 范数损失函数,也被称为最小平方误差损失函数(LSE),通过计算预测值与真实值间差值平方和的最小值获得,公式如下,其中 y_i,x_i 分别表示真实值和预测值。

$$L_2 = \sum_i^m (y_i - x_i)^2$$

如图 5 所示,横轴表示模型训练步数,纵轴表示模型准确率,当训练步数逐渐增加时,模型准确率也不断增加,当训练到 1 200 步左右时,模型准确率接近到 80%,其中最高准确率达到 82% 。与目前专利文本自动分类研究相比,准确率较高,说明本文提出的专利再分类模型能够有效进行过档文献再分类。如图 6 所示,横轴表示模型的训练步数,纵轴为损失函数值。随着训练步数增加,模型损失函数值不断降低,当模型训练至 1 000 步后,损失函数曲线趋于收敛,损失函数值小于 0.5,说明训练的模型已经逐渐接近最优状态,亦即模型已能达到最优准确率。综合模型准确率和损失函数曲线来看本文中训练的再分类模型已经取得较高的准确率,即使用该模型进行过档文献再分类能取得较好效果,分类准确率能达到 80% 左右,为后续使用再分类专利进行技术演化分析提供了数据质量保障。

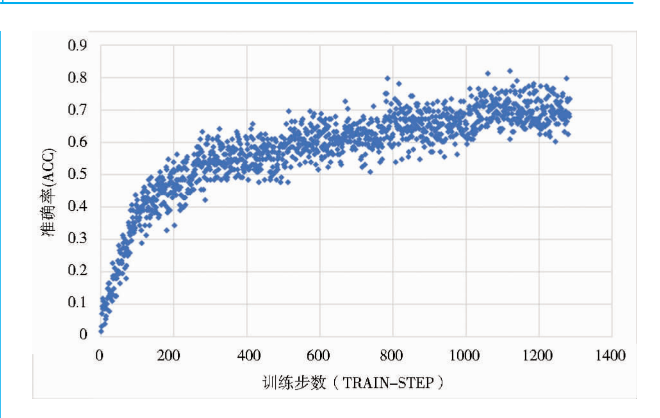


图 5 再分类模型的准确率曲线

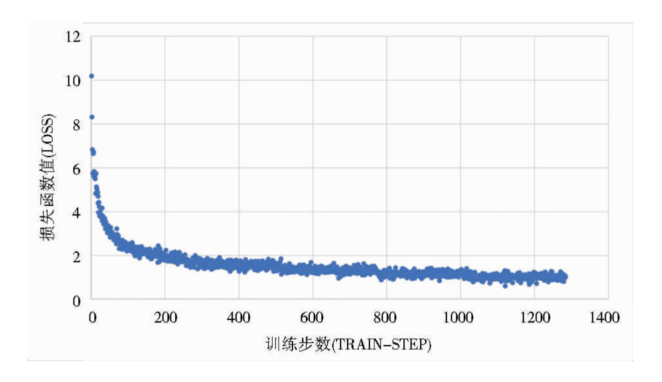


图 6 再分类模型的损失函数曲线

3 IPC 分类修订中的技术演化分析——以 H 部为例

3.1 申请趋势分析

首先分析 H 部专利公开趋势,如图 7 左图所示为 2009 年 – 2018 年 H 部、已修订和未修订分类的专利公开趋势。从图 7 来看,H 部分类的专利公开量逐年上升;H 部未修订分类的专利公开量在 2011 年 – 2014 年稍有增长;修订分类的专利公开量在 2009 年 – 2014 年平稳增长,但 2014 年后快速增长,并且修订分类的专利公开趋势线与 H 部专利公开趋势线走势大致相同,说明 H 部专利公开量变化受修订分类专利公开量变化的影响较大。从 H 部修订与未修订分类的专利公开量占比来看,如图 7 右图所示,2014 年之前修订分类与未修订分类的专利公开量在 H 部的全部专利公开中占比相对稳定,未修订分类的专利公开量稳定在 76% 左右,2014 年后未修订分类占 H 部专利分类公开量的比例逐年下降,到 2018 年已下降到约 60% ,而修订分类的专利公开量占比则从 2009 年的 24% 上涨到 2018 年的 40% ,也可见修订分类所表示技术发展迅速。

3.2 新增分类的技术演化分析

新增分类可以表示技术领域新发展,当新技术出现时,原有技术分类无法涵盖新技术,因此需要增加新分类来表示技术新进展。

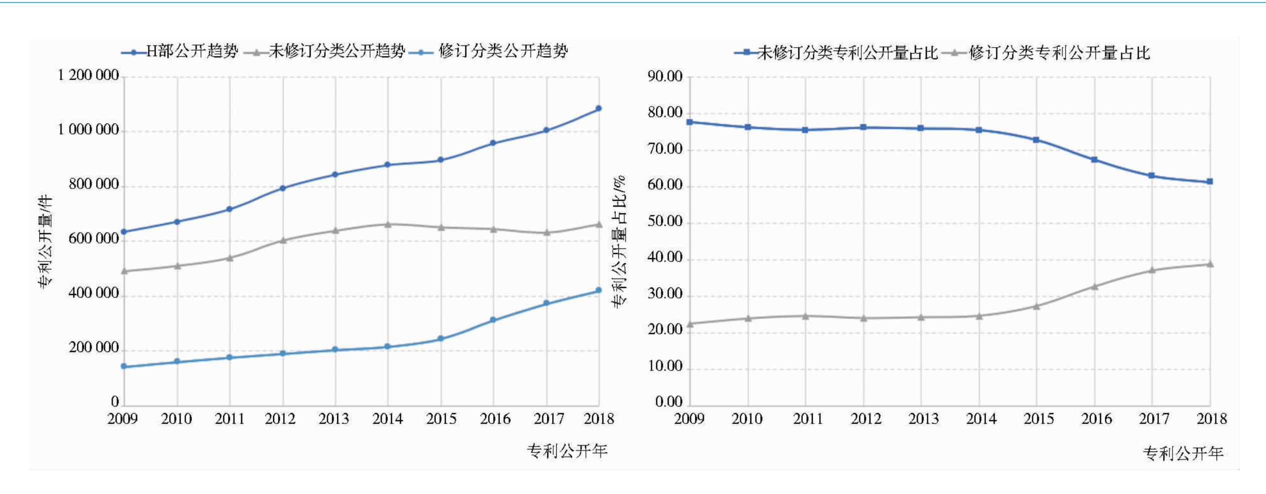


图 7　专利公开趋势（左图）和专利公开量占比（右图）

如表 3 所示，在 2009 年 –2018 年 10 年时间窗内，H 部的新增分类集中在 H01（基本电气元件）、H02（发电、变电或配电）、H04（电通信技术）3 个领域，涉及到 H01F（磁体；电感；变压器；磁性材料选择）、H01G（电容器；整流器、检波器、开关、光敏或热敏器件等）、H01L（半导体器件）、H01M（化学能转化为电能的装置，如电池组）、H01R（导电连接；相互绝缘电连接元件；集电器等）、H01Q（天线）、H02K（电机）、H02P（电动机、发电机或机电变换器的控制或调节；控制变压器、电抗器或扼流圈）、H03K（脉冲技术）、H04B（传输）、H04L（数字信息的传输，如电报通信）、H04N（图像通信，如电视）。从表 3 数据可知，2009 年 –2018 年 H 部新增分类中 H01L（半导体器件）和 H01M（化学能转化为电能的装置，如电池组）的专利公开量占比最大，分别为 24.78% 和 24.83%，说明该领域出现了较多技术创新；其次是 H04L（数字信息的传输，如电报通信）和 H04N（图像通信，如电视），分别占比 16.49% 和 17.36%，其余小类所代表技术的新增分类专利公开量都在 5% 及以下。

表 3　2009 年 –2018 年 H 部新增分类及其专利公开量

大类领域	小类	专利公开量（占比/%）	小类描述
H01 基本电气元件	H01F	2 731（0.29）	磁体；电感；变压器；磁性材料选择
	H01G	33 149（3.51）	电容器；整流器、检波器、开关、光敏或热敏器件等
	H01L	234 284（24.78）	半导体器件
	H01M	234 794（24.83）	化学能转化为电能的装置，如电池组
	H01R	13 785（1.46）	导电连接；相互绝缘电连接元件；集电器等
	H01Q	14 460（1.53）	天线
H02 发电、变电或配电	H02K	11 341（1.2）	电机
	H02P	18 973（2.01）	电动机、发电机或机电变换器的控制或调节；控制变压器、电抗器或扼流圈
H03 基本电子电路	H03K	661（0.07）	脉冲技术
H04 电通信技术	H04B	61 260（6.48）	传输
	H04L	155 976（16.49）	数字信息的传输，如电报通信
	H04N	164 197（17.36）	图像通信，如电视
合计		945 611（100）	

注：表中百分比为四舍五入后的数据

3.3　删除分类的技术演化分析

在技术发展过程中，既有新技术的产生也有老旧技术的淘汰和消亡，当旧分类方式不再适用于新技术发展时，会将旧分类从分类表中删除，因此删除分类可以表示被淘汰或消亡的技术。

如表 4 所示，2009 年 –2018 年间 H 部删除的分类主要涉及 H01（基本电气元件）、H02（发电、变电或配电）、H03（基本电子电路）、H04（电通信技术）4 个技术领域，包括 H01G（电容器；整流器、检波器、开关、光敏或热敏器件等）、H01H（电开关、继电器、选择器、紧急保护装置）、H01J（放电管或放电灯）、H01L（半导体器件）、H01M（化学能转化为电能的装置，如电池组）、

H01Q(天线)、H01R(导电连接;相互绝缘电连接元件;集电器等)、H02K(电机)、H02N(其他类目不包含的电机)、H03F(放大器)、H04B(传输)、H04J(多路复用通信)、H04L(数字信息的传输,如电报通信)、H04N(图像通信,如电视)、H04Q(选择)、H04W(无线通信网络)。其中 H04N(图像通信,如电视)、H04W(无线通信网络)、H04L(数字信息的传输,如电报通信)分别占比 26.11%、25.64%、21.38%是占比较大的几个类,说明这个几个领域被淘汰或消亡的技术较多,其他小类中除 H01L(半导体器件)的专利公开量占比 6%,剩余小类的专利占比皆在 5%以下。

表 4　2009 年 –2018 年删除分类及其专利公开量

大类领域	小类	专利公开量(占比/%)	小类描述
H01 基本电气元件	H01G	2 523(0.7)	电容器;整流器、检波器、开关、光敏或热敏器件等
	H01H	4(0.0)	电开关、继电器、选择器、紧急保护装置
	H01J	1 817(0.5)	放电管或放电灯
	H01L	21 726(6)	半导体器件
	H01M	11 870(3.28)	化学能转化为电能的装置,如电池组
	H01Q	2 136(0.59)	天线
	H01R	10 552(2.91)	导电连接;相互绝缘电连接元件;集电器等
H02 发电、变电或配电	H02K	798(0.22)	电机
	H02N	9 944(2.75)	其他类目不包含的电机
H03 基本电子电路	H03F	33(0.01)	放大器
H04 电通信技术	H04B	16 348(4.51)	传输
	H04J	1 130(0.31)	多路复用通信
	H04L	77 421(21.38)	数字信息的传输,如电报通信
	H04N	94 579(26.11)	图像通信,如电视
	H04Q	18 428(5.09)	选择
	H04W	92 865(25.64)	无线通信网络
合计		362 174(100)	

注:表中百分比为四舍五入后的数据

3.4　类内转移分类的技术演化分析

2009 年 –2018 年 H 部类内转移分类主要涉及 H01(基本电器元件)和 H04(电通信技术)中的分类,包括 2011 年修订的 H01R12/00 大组(用于印刷电路、扁平或带状电缆的导电连接)和 H01R24/00 大组(双部件连接的整体结构的导电连接);2012 年修订的 H01J11/00 大组(具有交流引发放电的充气放电管,例如 AC-PDP)以及 2018 年修订的 H04N13/00 大组(立体视频系统;多视点视频系统;其零部件)。

H01R12/00 大组(双部件连接的整体结构的导电连接)的修订和技术演化如图 8 和图 9 所示,H01R12/00 的修订是对其下位小组进行调整,增加了更细致的技术分支。图 9 中左侧表示修订前技术分类及专利量,右侧表示修订后技术分类和专利量,中间连线表示通过再分类使新旧版分类产生的联系(下同),可知 H01R12/71(用于刚性印刷电路或类似结构的耦合装置)是分类修订后的主要技术演化方向,约占 H01R12/00 大组 2009 年 –2010 年专利公开的 59.61%;其次为 H01R12/70(耦合装置),约占比 25.9%。

H01R24/00 大组(双部件连接的整体结构的导电连接)的修订是将旧版分类中的细分组全部删除,重新划分了新的细分组,修订前后的技术结构如图 10 所示,从图 11 中技术演化情况可知 H01R24/38(有同心或同轴布置的接触件)、H0R24/00(双部件连接的整体结构的导电连接)是分类修订后的主要技术演化方向,专利占比约 27.37% 和 25.33%。

H01J11/00 大组(立体视频系统;多视点视频系统;其零部件)修订前后的技术结构和技术演化如图 12 和图 13 所示。2012 年版 IPC 分类将 H01J11/02 和 H01J11/04 两个小组转移到 H01J11/10 – H01J11/54 小组,其中 H0J11/22(充气放电管电极,例如特殊形状、材料或结构)和 H0J11/46(有液池阴极的放电管在结构上与管子相连的一个或一个以上的电路零部件)是主要的技术演化方向,专利占比分别约 53.34% 和 17.72%。

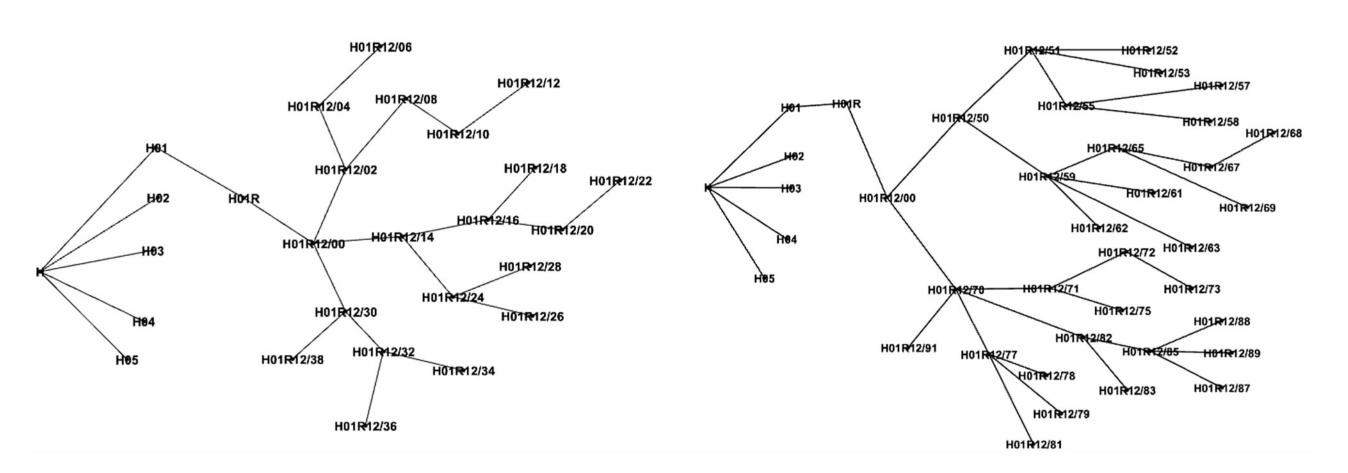


图 8　H01R12/00 大组修订前(左)和修订后(右)技术结构对比

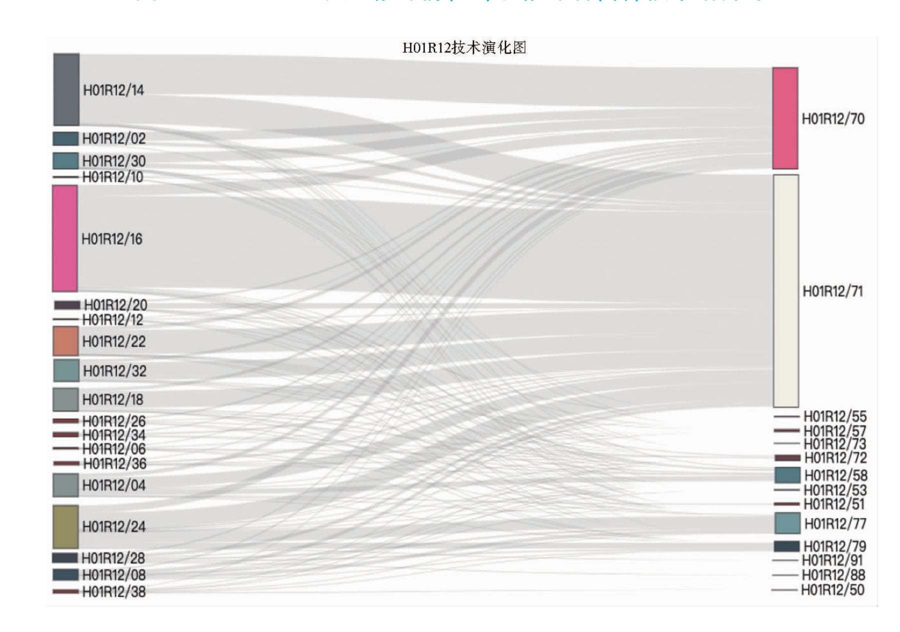


图 9　H01R12/00 大组的技术演化

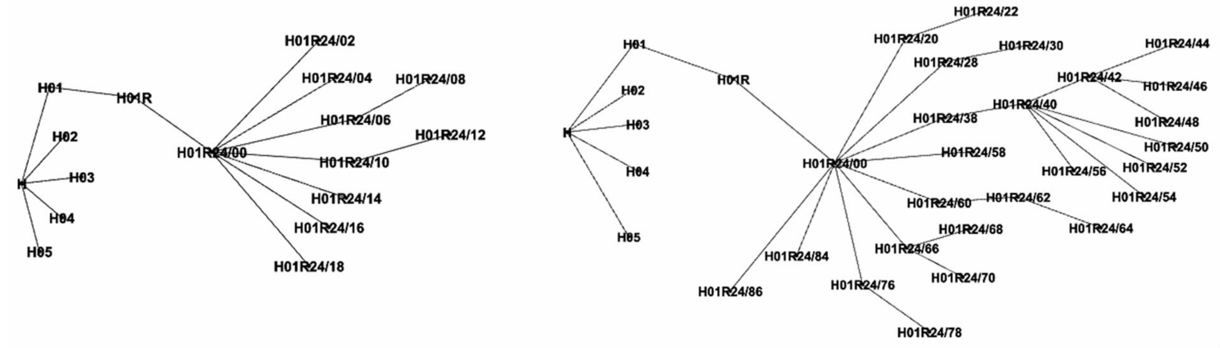


图 10　H01R24/00 大组修订前(左)和修订后(右)技术结构对比

H04N13/00 大组(立体视频系统;多视点视频系统;其零部件)修订前后的技术结构和技术演化如图 14 和图 15 所示。2018 年版 IPC 分类表将 H04N13/02 和 H04N13/04 重新分类到 H04N13/10－H04N13/198 小组，技术演化结果显示 H04N13/30(图像重现装置)、H04N13/20(图像信号发生器)、H04N13/10(立体或者多视点图像信号的处理,记录或传输)是主要的技术演化方向,专利占比分别约为 25.92%、13.03% 和 10.13%。

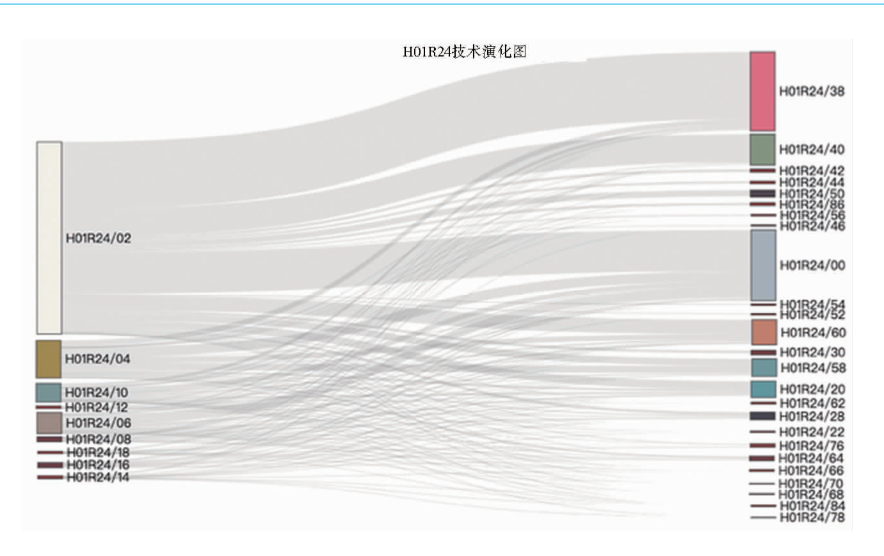


图 11 H01R24/00 大组的技术演化

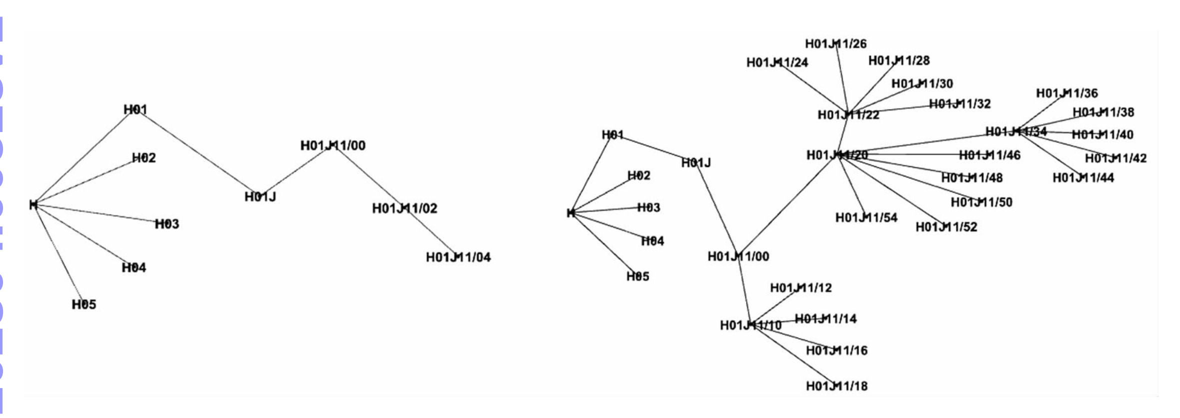


图 12 H01J11/00 大组修订前(左)和修订后(右)技术结构对比



图 13 H01J11/00 大组的技术演化

3.5 类间转移分类的技术演化分析

2009 年 – 2018 年 H 部类间转移分类主要涉及 H02(发电、变电或配电)和 H04(电通信技术)中的分类,包括 2009 年修订的 H04W(无线通信网络);2011 年修订的 H04N21/00(可选的内容分发,例如交互式电视,或视频点播);2014 年修订的 H02S(由红外线辐射、可见光或紫外光转换产生电能,如使用光伏(PV)模块)。H04W 小类是 2009 年将 H04Q7/00 大组向 H04W 小类的转移,修订前后的技术结构和技术演化如图 16 和图 17 所示,其中 H04W4/00(专门适用于无

线通信网络的业务;其设施)、H04W4/02(利用位置信息的无线通信网络业务)、H04W4/12(消息传送;邮箱;通告)、H04W4/16(与通信相关的补充业务,例如,呼叫转移或呼叫保持)、H04W4/06(广播选择分发业务,例如,多媒体广播组播业务〔MBMS〕;到用户组的业务;单向选呼业务)等是主要的技术演化方向,专利占比分别约为 48.16%、10.44%、10.54%、7.47%、6.65%。

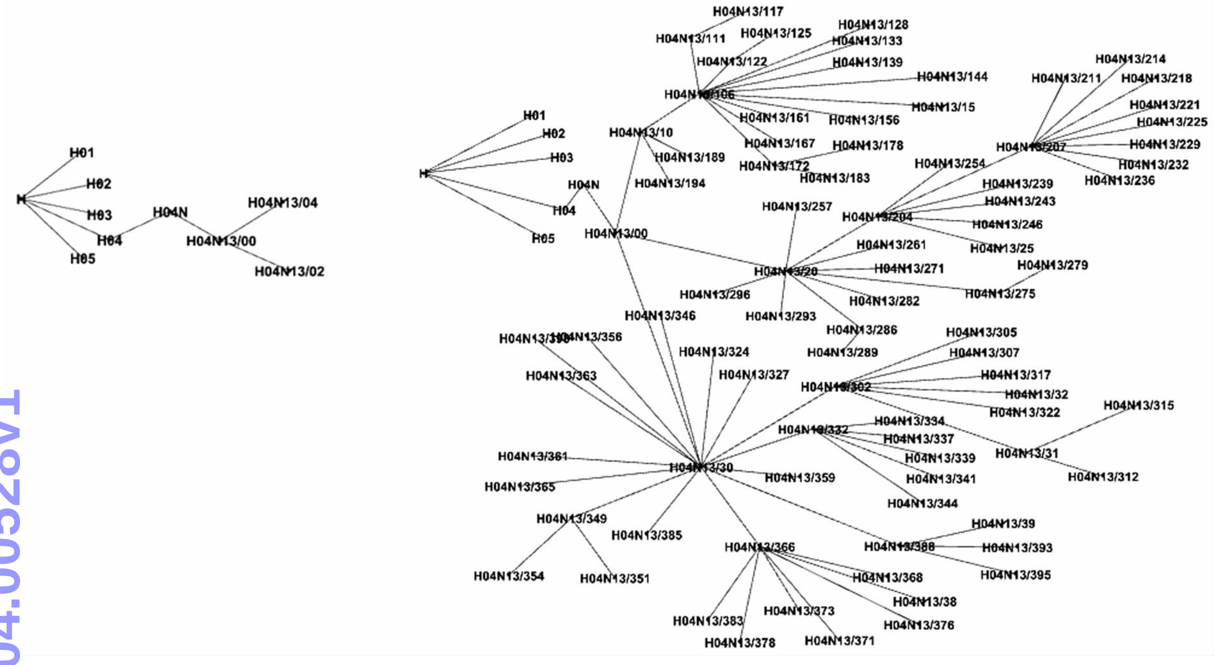


图 14 H04N13/00 大组修订前(左)和修订后(右)技术结构对比

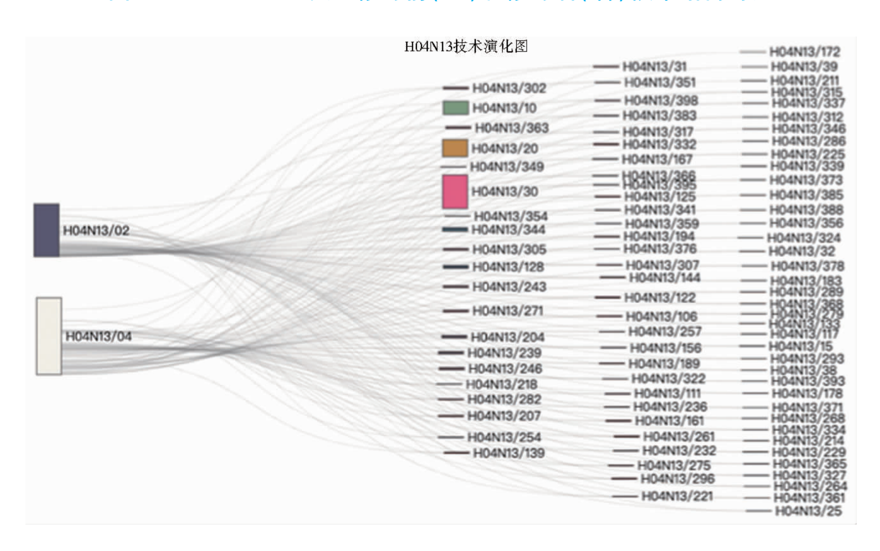


图 15 H04N13/00 大组的技术演化

H04N21/00 大组的修订是将 H04N7/58、H04N7/60 和 H04N7/62 转移到 H04N21/00 大组,修订前后技术结构和技术演化如图 18 和图 19 所示。其中 H04N21/43(个人计算机内容接收或内容交互)、H04N21/242(同步进程,例如程序时钟调用的处理)是主要的技术演化方向,分别占比 13.12% 和 10.0%。

H02S 小类的修订是将 H02N6/00 和 H01L31/045 和 H01L31/058 三个小组合并并转移到 H02S 小类,修订前后的技术结构和技术演化如图 20 和图 21 所示。其中 H02S20/32(专门用于太阳能跟踪的光伏模块支撑结构)、H02S40/22(光伏模块的反光或集光的设备)、H02S30/20(可折叠的光伏模块)是主要技术演化方向,分别占比约 12.47%、9.74% 和 6.73%。

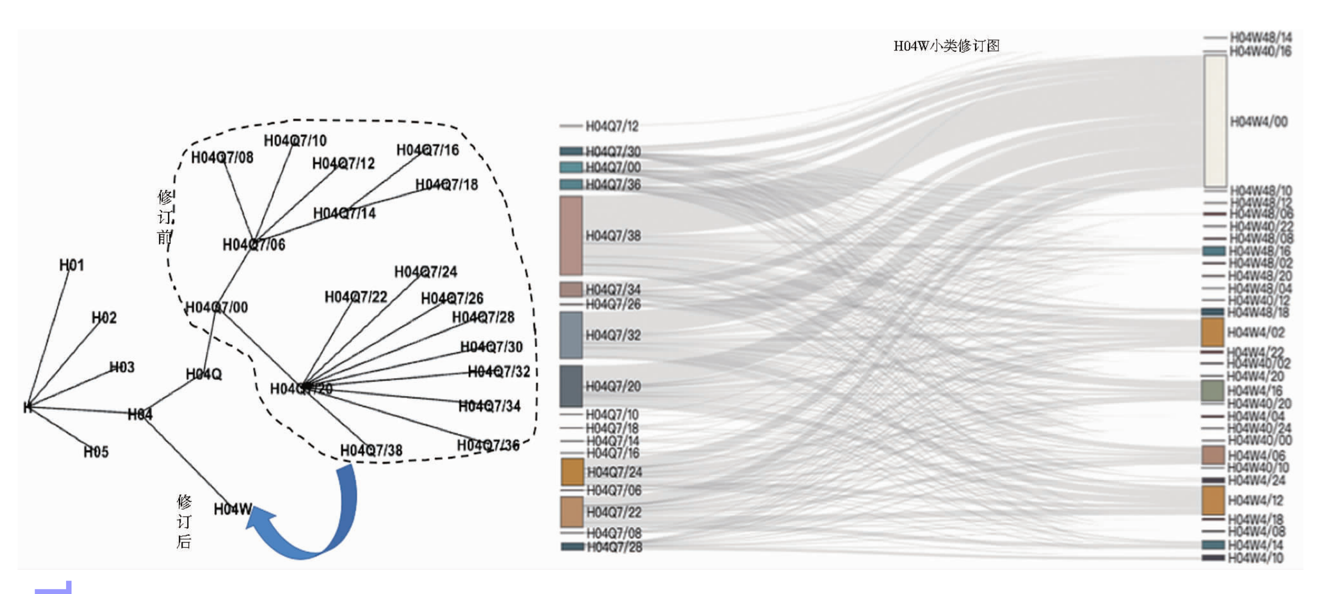


图 16 H04W 小类修订前后的技术结构对比

图 17 H04W 小类的技术演化

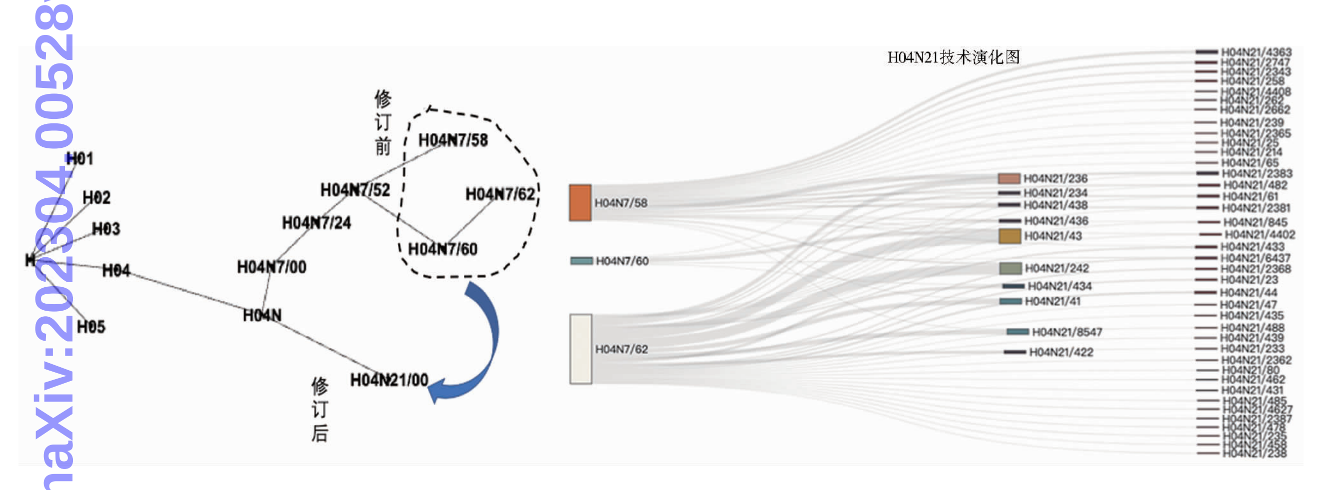


图 18 H04N21/00 大组修订前后的技术结构对比

图 19 H04N21/00 大组技术演化

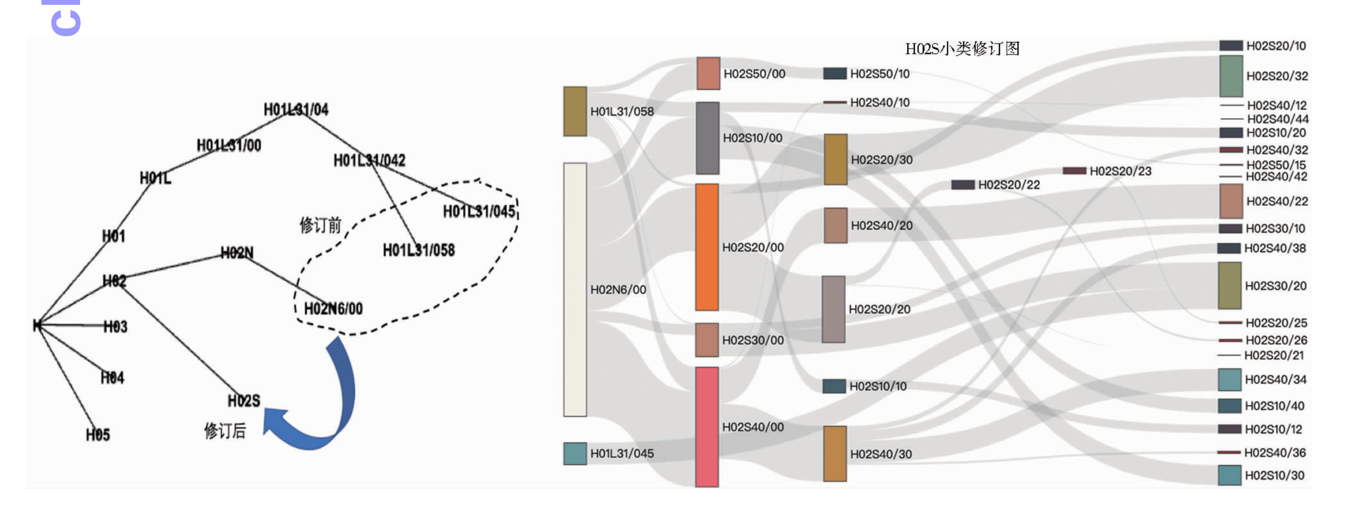


图 20 H02S 小类修订前后的技术结构对比

图 21 H02S 小类技术演化

4 结语

技术演化已有完整的研究体系,但从专利分类修订角度开展技术演化的研究则较少,本文从这个角度开展初步探索。首先从专利分类修订出发,根据2009年 -2018 年 IPC 分类表 H 部的修订总结出 4 种分类修订类型:新增分类、删除分类、类内转移分类、类间转移分类,并提出分类修订后产生过档文献这一概念。其次针对过档文献导致新旧版专利分类表无法衔接问题提出基于 Word2vec + TextCNN 模型的专利再分类方法,对过档文献再分类,从而通过再分类专利联系新旧版专利分类表。最后结合分类修订前后技术结构变化和对再分类专利的统计分析,描述分类修订中主要技术演化方向。本文不仅能为技术演化研究提供新视角,扩展技术演化研究体系,还能为过档文献的再分类工作提供参考,是一次有益探索。

本文研究还存在以下不足:首先未能对 H 部技术演化情况作出全景描述,未能通过可视化手段实现"技术演化全景图"展示,而是分别对 2009 年 -2018 年修订分类进行解读,缺乏全局性和整体性。其次,Word2vec + TextCNN 模型准确率目前仅能达到 80% 左右,再分类准确率有待提高,这对本研究中提供技术演化情报的准确性影响巨大。最后未能提出描述技术演化的计量指标,只采用比例形式描述技术演化重点方向,难以深度揭示技术演化信息,未来可进一步完善该研究。

参考文献:

[1] 谢寿峰. 基于专利分析的技术演变与预测研究[D]. 南京:南京理工大学,2014.
[2] 刘云, 刘璐, 闫哲. 基于专利计量的全球碳纳米管领域技术创新特征分析[J]. 科研管理, 2016(S1):337 -345.
[3] SUZUKI J, KODAMA F, et al. Technological diversity of persistent innovators in Japan: two case studies of large Japanese firms [J]. Research policy, 2004, 33(3):531 -549.
[4] 廖列法, 勒孚刚. 基于 LDA 模型和分类号的专利技术演化研究[J]. 现代情报, 2017(5):15 -20.
[5] 陈伟,林超然,李金秋,等. 基于 LDA-HMM 的专利技术主题演化趋势分析——以船用柴油机技术为例[J]. 情报学报,2018,37(7):732 -741.
[6] YOON J, KIM K. Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks[J]. Scientometrics, 2011, 88(1):213 -228.
[7] SASAKI H, KAJIKAWA Y, SAKATA I, et al. Predicting the potential industrial fields of technological spin-offs by using IPC in patent analysis[C]// Proceedings of technology management for emerging technologies. Vancouver:IEEE, 2012:975 -991.
[8] ZHENG J, CHEN D Z, HUANG M H, et al. Industry and technology development in China from 2003 to 2008: a perspective from patent classification analysis [C]//Technology management for global economic growth. Phuket:IEEE, 2010:2256 -2263.
[9] ZHOU X, ZHANG Y, PORTER A L, et al. A patent analysis method to trace technology evolutionary pathways[J]. Scientometrics, 2014, 100(3):705 -721.
[10] JUN S. IPC code analysis of patent documents using association rules and maps-patent analysis of database technology[C]//International conference on database theory and application/international conference on bio-science and bio-technology. Jeju Island: SPRINGER-VERLAG BERLIN,2011:21 -30.
[11] 黄斌, 黄鲁成, 吴菲菲, 等. 基于专利共类的技术关联特征识别[J]. 情报杂志, 2015(1):44 -48.
[12] CHEN J K C, PHAM V K, LIN F C, et al. Studying the patent of technology development on Dye Sensitized Solar Cell[C]/ Technology management in the it-driven services. San Jose: IEEE, 2012: 985 -996.
[13] 李瑞茜, 陈向东. 基于专利共类的关键技术识别及技术发展模式研究[J]. 情报学报, 2018, 37(5):49 -56.
[14] KRAFFT J, QUATRARO F, SAVIOTTI P P. The knowledge-base evolution in biotechnology: a social network analysis[J]. Economics of innovation and new technology, 2011, 20(5):445 -475.
[15] 方曙,胡正银,庞弘燊,等. 基于专利文献的技术演化分析方法研究[J]. 图书情报工作,2011,55(22):42 -46.
[16] 吴红, 伊惠芳, 马永新, 等. 面向专利技术主题分析的 WI-LDA 模型研究[J]. 图书情报工作, 2018,62(17):68 -74.
[17] 雷滔,陈向东. ICT 领域跨学科关联度的块段模型分析——基于我国校企合作申请的专利[J]. 中国软科学,2011(1):67 -74.
[18] 苗红,王晓宇,黄鲁成,等. 基于专利的跨领域技术特征测度与分析[J]. 情报学报,2015,34(5):484 -492.
[19] LAFOND F, KIM D. Long-run dynamics of the U. S. patent classification system [J]. Journal of evolutionary economics, 201, 29 (2):631 -664.
[20] EUSEBI C A, SILBERGLITT R. Identification and analysis of technology emergence using patent classification[R]. California: RAND Corporation, 2014.
[21] WANG C C, SUNG H Y, HUANG M H. Technological evolution seen from the USPC reclassifications[J]. Scientometrics, 2016, 107(2):537 -553.
[22] 王文晶. 新能源汽车产业专利分类修订与产业发展关联初探[J]. 科技创新与应用, 2017(14):25 -26.
[23] 牛力, 胥凯, 王扬平. 空调领域 IPC 的修订及在检索中的应用[J]. 中国发明与专利, 2012(S1):136 -139.
[24] 世界知识产权组织(WIPO) [EB/OL]. [2020 -11 -14].. http://www.wipo.int/classifications/ipc/zh/.
[25] 世界知识产权组织. Guidelines for revision of the IPC [EB/OL].

[2020 - 11 - 14]. https://www.wipo.int/export/sites/www/classifications/ipc/en/general/guidelines_revision_ipc.pdf.
[26] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[J]. arXiv preprint. arXiv:1301.3781, 2013.
[27] KIM Y. Convolutional neural networks for sentence classification [C]// Proceedings of the 2014 conference on empirical methods in natural language processing. Doha: Association for Computational Linguistics, 2014:1746 - 1751.

作者贡献说明:
许景龙:结构构思、研究实施、论文撰写和修改;
赵亚娟:论文选题、思路把握、论文修改与审定。

Research on Technology Evolution in the Revised IPC Classification

Xu Jinglong[1] Zhao Yajuan[2,3]
[1] Library of Party School of the Guangdong Provincial Committee of CPC (Guangdong Institute of Public Administration), Guangzhou 515300
[2] National Science Library, Chinese Academy of Sciences, Beijing 100190
[3] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

Abstract: [**Purpose/significance**] This article attempts to study technological evolution from the perspective of patent classification revisions, and provides new ideas for technological evolution research. [**Method/process**] First, according to the 2009 - 2018 revisions of the H part of the IPC classification table, four types of revisions were summarized: new classification, deletion classification, intra-class transfer classification, and inter-class transfer classification. Secondly, for the archived documents generated after classification revision, a reclassification method of patent archived documents based on the Word2vec + TextCNN model was proposed, so that the old and new classification tables were connected by reclassifying patents. Finally, combined with the revised classification and reclassification patents of Part H from 2009 to 2018, the preliminary exploration of technological evolution was carried out. [**Result/conclusion**] The patent reclassification model can effectively solve the problem of archived documents, provides reference for patent reclassification work, and can link the new and old patent classification tables; based on IPC classification revision and reclassification of patents, the main technological evolution direction after classification revision can be analyzed, providing a new perspective for research on technological evolution.

Keywords: IPC classification patent classification revision patent reclassification technology evolution